Atmospheric Pollution Research
www.atmospolres.com

Assessment of WRF/Chem $PM_{2.5}$ forecasts using mobile and fixed location data from the Fairbanks, Alaska winter 2008/09 field campaign

Nicole Mölders [1], Huy N.Q. Tran [1], Catherine F. Cahill [2], Ketsiri Leelasakultum [1], Trang T. Tran [1, 3]

[1] *University of Alaska Fairbanks, Geophysical Institute, College of Natural Science and Mathematics, Department of Atmospheric Sciences, 903 Koyukuk Dr., Fairbanks, AK 99775, USA*
[2] *University of Alaska Fairbanks, Geophysical Institute, College of Natural Science and Mathematics, Department of Chemistry, 903 Koyukuk Dr., Fairbanks, AK 99775, USA*
[3] *University of Alaska Fairbanks, College of Natural Science and Mathematics, Department of Chemistry, 903 Koyukuk Dr., Fairbanks, AK 99775, USA*

ABSTRACT

Weather Research and Forecasting model inline coupled with a chemistry package $PM_{2.5}$ forecasts were assessed using fixed–site $PM_{2.5}$ concentration and specification, and mobile $PM_{2.5}$ concentration and temperature measurements from the Fairbanks winter 2008/09 field campaign. Performance differs with concentrations, varies among months and sites, and best results are achieved for $PM_{2.5}$ concentrations between 15 and 50 μg/m^3. On average over half–a–year and all sites, 24 h–average $PM_{2.5}$ concentrations have a fractional bias and error, and a normalized mean bias and error of 22%, 67%, 13% and 71%, respectively. The skill scores derived from the mobile measurements indicate that high data density increases the representativeness of the observations and enhances the evaluation of spatial details. The model performed well for organic carbon and acceptably for sulfate, but underestimated ammonium significantly.

$PM_{2.5}$ concentrations measured by two different devices at the same site indicate that measurement errors at extremely low temperatures and humidities explain up to 24% of the normalized mean error. Some discrepancies can be attributed clearly to errors in emissions, chemical boundary conditions and meteorology.

Keywords:
WRF/Chem
High latitudes
Speciation
Mobile measurements
$PM_{2.5}$

Article History:
Received: 03 August 2011
Revised: 25 November 2011
Accepted: 27 November 2011

Corresponding Author:
Nicole Mölders
Tel: +1-907-474-7910
Fax: +1-907-474-7379
E-mail: molders@gi.alaska.edu

doi: 10.5094/APR.2012.018

1. Introduction

In the past, photochemical air quality models (AQMs) were not evaluated for use at high latitudes as most air quality issues were related to ozone or particulate matter of diameter smaller than 2.5 μm ($PM_{2.5}$) in low or mid–latitudes. An assessment of AQM performance for high latitudes became necessary when Fairbanks, Alaska was designated a $PM_{2.5}$ nonattainment area (NAA) after the tightening of the 24 h–average National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$ to 35 μg/m^3 in 2006.

The Fairbanks' nonattainment issue is a local one. Observations combined with HYSPLIT trajectories (Draxler et al., 2009) and photochemical modeling show that the region receives only minor amounts of pollution from long–range transport (Cahill, 2003; Tran et al., 2011). The major sources of primary particulate matter are within the NAA. Typically, $PM_{2.5}$ exceedances occur during strong temperature inversions on calm wind days when the inversion traps local emissions from heating and vehicles near the surface (Tran and Mölders, 2011). Past speciation data indicated that secondary aerosol components constitute about 36% of the $PM_{2.5}$ mass. In ranked order, the most important $PM_{2.5}$ components are organic carbon (OC), sulfate (SO_4^{2-}), elemental carbon (EC), nitrate (NO_3^-), and ammonium (NH_4^+).

Mölders et al. (2011) assessed the performance of the Weather Research and Forecasting model (Skamarock et al., 2008) inline coupled with a chemistry package (WRF/Chem; Peckham et al., 2009) in simulating subarctic boundary layer characteristics of winter 2005/06. They also used data from four aerosol sites, of which two had $PM_{2.5}$ data. They found a strong relation between errors in $PM_{2.5}$ concentrations and temperature errors. Difficulty in simulating the temporal evolution of aerosol concentrations occurred when WRF/Chem mistimed frontal passages or missed to capture sudden temperature changes or the full inversion strength. WRF/Chem largely underestimated NO_3^- at the three remote sites and $PM_{2.5}$ at the polluted site (Fairbanks).

The robustness of any operational evaluation depends on the amount and quality of observations; extensive data from field campaigns provide the best basis for assessing AQM performance (e.g. Djalalova et al., 2010). Until 2008, the State Office Building (SB) was the only $PM_{2.5}$ monitoring site in Fairbanks. In winter 2008/09, the Fairbanks North Star Borough supported a field campaign to assess the situation in the NAA. This dataset provides a first time opportunity to evaluate WRF/Chem for high latitudes over an entire winter. The scope of our study was to analyze WRF/Chem's ability to simulate $PM_{2.5}$ concentration using this data and to assess the suitability of mobile measurements for AQM evaluation.

2. Experimental Design

2.1. Simulations

We used the Alaska–adapted WRF/Chem setup as described in Mölders et al. (2011). This means the WRF–Single–Moment cloud–microphysics scheme (Hong and Lim, 2006), the 3D version of Grell and Devenyi's (2002) cumulus–ensemble approach, the Goddard two–stream, multi–band model, the Rapid Radiative Transfer Model (Mlawer et al., 1997), Janjic's (2002) atmospheric boundary layer and sublayer–schemes, and a modified version of Smirnova et al.'s (2000) land–surface model. Furthermore, Stockwell et al.'s (1990) gas–phase chemical mechanism, Madronich's (1987) photolysis rates calculation, Wesely's (1989) deposition module with the modifications introduced by Mölders et al. (2011), and the Modal Aerosol Dynamics Model for Europe (MADE; Ackermann et al., 1998) and Secondary Organic Aerosol Model (SORGAM; Schell et al., 2001) were used.

The model domain centered over Fairbanks covered Interior Alaska (Figure 1) with a horizontal grid increment of 4 km and a vertically stretched grid to 100 hPa. Analysis was performed on a domain of 80×70 grid points.

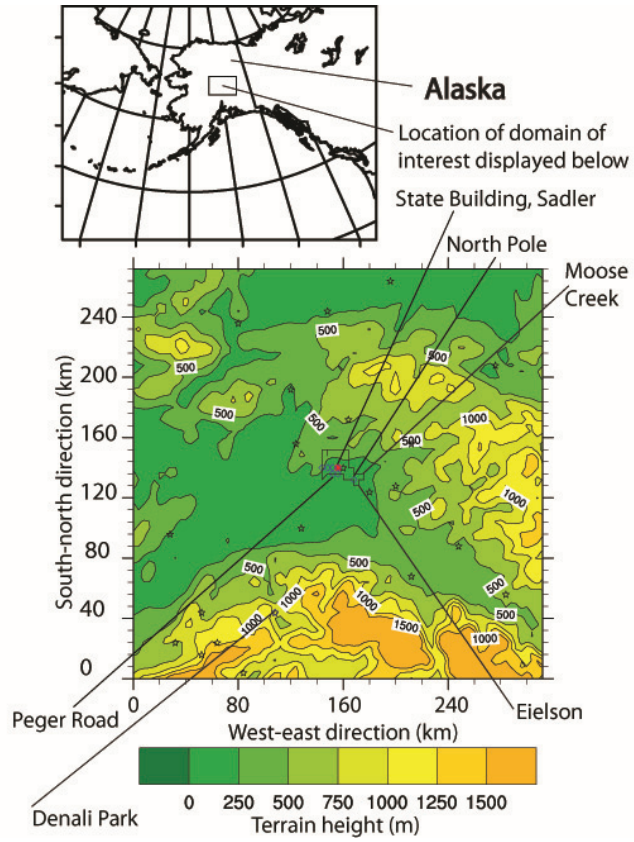

Figure 1. *Location of the interest area and topography therein. Stars, diamonds, and the dot indicate the surface meteorological sites, $PM_{2.5}$ sites and a MET–tower. The polygon marks the NAA.*

Anthropogenic emissions were based on the National Emission Inventory (NEI) of 2008. As the NEI2008 had no point source emissions for the domain at the time of performing the simulation, information, such as emissions provided by the facility operators, was used. Otherwise, we assumed a 1.5%/y increase from the last NEI. Area and line emissions were allocated in space and time depending on relevant data like population density, traffic counts, land–cover, month, weekday, hour, and emission source types. For emissions from traffic, power generation and heating, a temperature dependency was considered that leads to higher (lower) emissions for temperatures below (above) the 1971–2000 mean. Plume–rise was calculated for point emissions following Peckham et al. (2009). Biogenic emissions were calculated following Simpson et al. (1995).

The initial conditions for the meteorological, snow and soil quantities were downscaled from the 1°×1°, 6 h–resolution National Centers for Environmental Prediction global final analyses. This dataset also served to downscale and provide downscaled meteorological boundary conditions. Idealized vertical profiles of Alaska background concentrations for the chemical species served to initialize the chemical fields. Since Fairbanks is far remote from any emission sources, Alaska background concentrations served as chemical boundary conditions.

We ran WRF/Chem in forecast–mode for 10–01–2008 to 04–01–2009 (called OTM hereafter). The chemical distributions obtained at the end of a simulation served as chemical initial conditions for the next simulation, while the meteorology was initialized every five days.

2.2. Observations

The borough made hourly observations of total $PM_{2.5}$ mass using Met–One Beta Attenuation Monitors (BAM 1020) at the SB and Peger Road (PR) for OTM, and a Regional Air Monitoring Systems (RAMS) at two different locations (called RAMS1 and RAMS2) for October 14 to November 3 and November 7 to December 5. They moved the RAMS to three other locations during OTM, but had technical issues. Thus, we excluded this data from the analysis. We determined 24 h–averages from the hourly data as the 24 h–average is relevant for the NAAQS.

Filter based 24 h–average 1–in–3–days $PM_{2.5}$ concentrations obtained with the Federal Reference Method (FRM) exist for the SB, PR, North Pole (NP), and Sadler sites. Speciation data for species contributing to total $PM_{2.5}$ mass collected from filter based 24 h–average concentrations by Met–One Super SASS Speciation Monitors of the Speciation Trends Network (STN) exist every 1–in–3–days at the SB for OTM and at PR and NP for January to March. The SB, NP and PR sites are located in downtown Fairbanks, a mixed commercial–industrial–residential area, and an industrial area, respectively. The Denali Park (DP) Interagency Monitoring of Protected Visual Environments (IMPROVE) site is the only remote site and only site outside the NAA (Figure 1). It has 1–in–3–days $PM_{2.5}$ and speciation data.

The borough took mobile measurements of $PM_{2.5}$ concen–trations and temperatures using vehicles instrumented with a BGI $PM_{2.5}$ sharp–cut–cyclone, sample–liner heaters, Garmin GPS, drycal flow calibrator, and temperature loggers. The vehicles drove along predetermined routes in the NAA on 15, 22, 24, 13, and 12 days in November, December, January, February, and March, respectively, providing 664 000 data points (1 every 2 s).

We performed a quality assurance/quality control (QA/QC) that discarded all temperature and $PM_{2.5}$ data for which the measured temperature deviated more than the 1971–2000 monthly–mean diurnal temperature range from the mean temperature determined from all temperature data of the respective drive. This QA/QC served to discard data taken when the vehicle pulled out and the sensors were still adjusting to the outside air. Since occasionally plumes from trucks or buses that emit at about the sniffer height (~2.44 m), may have hit the sniffer, the QA/QC procedure discarded all $PM_{2.5}$ concentrations that differed >5 $\mu g/m^3$ between two consecutive measurements. We projected the remaining data onto the model grid and averaged over all measurements that were taken in the same grid cell and hour. We compared these hourly–spatially–averaged observations to the hourly volume averages obtained from WRF/Chem.

Meteorological surface observations at 23 sites (Figure 1) of hourly 2 m–temperature, 2 m–dewpoint temperature, and 10 m–wind speed were available from the Western Region Climate Center. At two sites, sea level pressure (SLP), and at 20 sites, 24 h–accumulated solar radiation were available. Starting January 17 2009 0200UTC hourly temperatures at 3 m, 11 m and 22 m, wind speed and direction at 11 m, relative humidity at 3 m and shortwave downward radiation were available from a meteorological tower (MET–tower) located downtown.

2.3. Analysis

We determined the root mean square error (RMSE), bias (simulated – observed), correlation, and standard deviation of the meteorological quantities. For the chemical quantities we calculated the fraction of simulated concentrations (C_s) being within a factor of two of the measured concentrations (C_o), the mean fractional bias,

$$FB = \frac{2}{N}\sum_{i=1}^{N}\left(C_{s,i} - C_{o,i}\right)/\left(C_{s,i} + C_{o,i}\right) \times 100\% \quad (1)$$

fractional error,

$$FE = \frac{2}{N}\sum_{i=1}^{N}\left|C_{s,i} - C_{o,i}\right|/\left(C_{s,i} + C_{o,i}\right) \times 100\% \quad (2)$$

normalized mean bias,

$$NMB = \sum_{i=1}^{N}\left(C_{s,i} - C_{o,i}\right)\Big/\sum_{i=1}^{N} C_{o,i} \times 100\% \quad (3)$$

and normalized mean error,

$$NME = \sum_{i=1}^{N}\left|C_{s,i} - C_{o,i}\right|\Big/\sum_{i=1}^{N} C_{o,i} \times 100\% \quad (4)$$

Since the acceptable range for bias and error is larger at low than high concentrations, we used Boylan and Russell's (2006) performance goals and criteria.

For fixed sites we (1) extracted the data simulated for the grid cell the sites fell into and (2) alternatively determined the "value at the site" by a distance weighted closest neighboring averaging of the simulated values around the site's locations. Since the values insignificantly differed for all sites with a correlation of 0.999, we presented the evaluation results using data extracted with method (1).

3. Results

3.1 MET–tower

The MET–tower data revealed that WRF/Chem had a small temperature bias of 0.6 K, 0.7 K and 1.1 K at 3, 11 and 22 m, respectively. Correlation between simulated and observed temperature was 0.880, 0.886 and 0.890 at these heights. The bias of wind direction at 11 and 22 m, relative humidity, and SLP were –47°, –55°, 16%, and –0.2 hPa, respectively. The approximately 5 W/m^2 shortwave radiation bias is within the margin of observational accuracy.

The MET–tower data suggest that WRF/Chem overestimated vertical mixing, which may have over–diluted concentrations. On average, WRF/Chem simulated temperature gradients of 1.08 K/100 m and 0.09 K/100 m, while 0 K/100 m and 0.01 K/100 m were observed between 3 and 11 m and 11 and 22 m, respectively. WRF/Chem had positive wind speed biases of 1.15 m/s (2.39 m/s) at 11 m (22 m) with wind speed RMSEs of 1.84 m/s (3 m/s). At 11 m (22 m), the mean simulated and observed wind speeds were 1.97 m/s (3.46 m/s) and 0.82 m/s (1.07 m/s), respectively.

3.2. Meteorological surface observations

On average over all 23 surface meteorological sites, WRF/Chem captured the temporal evolution of the meteorological quantities well (Figure 2). Discrepancies between simulated and observed meteorological quantities occurred due to mistiming of frontal passages and after sudden strong temperature changes. On average over OTM and all 23 sites, the temperature, dewpoint temperature, SLP, wind speed and direction biases were 1.3 K, 2.1 K, –1.9 hPa, 1.55m/s, and –4°, respectively.

WRF/Chem performed best for SLP, followed by dewpoint and air temperature. Air and dewpoint temperature scores varied the most among months indicating WRF/Chem's difficulties in modeling the frequent inversions. On average over all 23 sites, the monthly–averaged observed temperatures were –9.7 °C, –15.4 °C, –19.9 °C, –21.9 °C, –16.3 °C, and –15.2 °C in October, November, December, January, February and March, respectively. WRF/Chem overestimated the monthly average temperatures leading to biases of 0.5 K, 0.8 K, 2 K, 2.6 K, 1.6 K and 0.3 K and RMSEs of 3.8 K, 4.8 K, 6.1 K, 4.3 K, 5.2 K and 4.1 K in OTM. The dewpoint temperature RMSEs were around 4 K in all months except November (6.2 K) and December (4.8 K). The correlation coefficient of simulated and observed air (dewpoint) temperature was 0.897 (0.905) for OTM. WRF/Chem captured the air and dewpoint temperature variance best in October and March (Figure 2). It underestimated the frequency of temperatures greater than –5 °C and lower than –30 °C, and failed to reproduce temperatures lower than –40 °C.

The errors in air and dewpoint temperature led to a RMSE of 15% and wet bias of 5% in relative humidity over OTM. Correlation of simulated and observed relative humidity varied among months and was 0.473 over OTM. Relative humidity forecasts were best for January, but WRF/Chem underestimated strongly the variance. Due to the air and dewpoint temperature errors over OTM the simulated frequency of relative humidity less than 60% was slightly underestimated. These relative humidity errors may affect aerosol formation.

WRF/Chem overestimated the variance and magnitude of wind speed leading to correlations less than 0.6 in all months and 0.573 over OTM (Figure 2). It failed to capture the high frequency of calm winds (<1 m/s), but forecasted wind speed and its variance best for January, the calmest month. WRF/Chem had difficulties in predicting wind direction, but acceptably captured its variance in all months. It provided a slightly too high (low) probability of wind direction between 30° and 100° (140° and 220°) due to terrain differences between the model and the real world. WRF/Chem like other AQMs uses the mean terrain height within a grid cell and ignores subgrid scale terrain heterogeneity that is inherent to the observations.

It is well known that models have difficulty simulating calm winds accurately. Zhao et al. (2011) reported that WRF had difficulty reproducing weak surface winds (<1.5 m/s) in their long–term 4 km–increment simulation over California. This difficulty led to bias (RMSE) of more than 3 m/s (4.5 m/s). As in our study, their weak winds were associated with stagnation events. Our bias (RMSE) was 1.55 m/s (2.4 m/s). Thus, our wind forecasts have to be considered as good under the given situation.

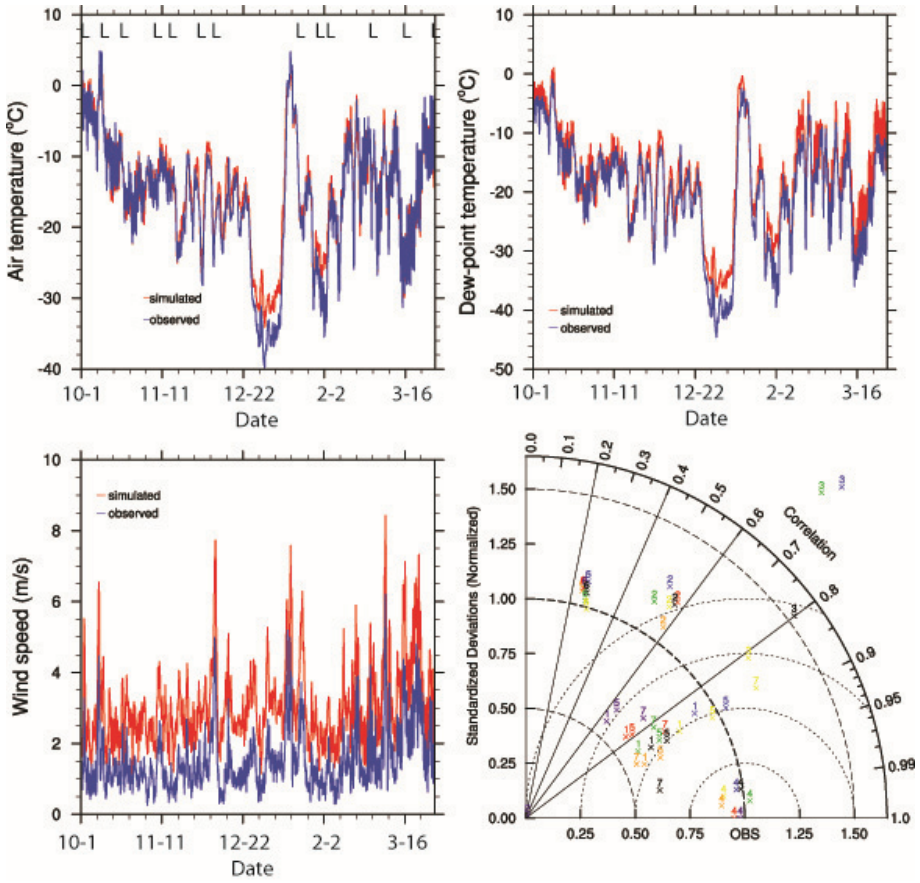

Figure 2. *Time series of various simulated and observed meteorological quantities averaged over all sites for which data were available, and Taylor diagram displaying performance statistics. L indicates times of frontal passages. In the Taylor diagram, the numbers 1 to 7 represent the scores for air temperature, wind speed, downward shortwave radiation, SLP, dewpoint temperature, and wind direction from all available surface meteorological data and the mobile measured temperature for winter 2008/09 (black), October (blue), November (purple), December (red), January (orange), February (green) and March (yellow). The solid, dotted and dashed lines are correlation, normalized RMSE (increment of 2), and normalized standard deviation. OBS indicates a perfect forecast. Scores on the bold–dashed arc have correct standard deviations.*

3.3. Mobile temperature measurements

The average temperatures over all drives were –16.6 °C, –23.6 °C, –27.8 °C, –19.2 °C, and –11.8 °C in November, December, January, February and March, respectively. Along the drive routes, WRF/Chem had temperature biases of 0.1 K, 0.8 K, 1.7 K, 0.3 K, and –0.1 K, and RMSEs of 3.9 K, 5.3 K, 7.4 K, 5.7 K, and 3.9 K from November to March, respectively. The warm bias was largest in the coldest month (January), while RMSEs and biases were lowest in the warmest months (November, March). Despite this behavior and the highest temperature variance in January, simulated and observed temperatures correlated better in January than in November (Figure 2).

The RMSEs (biases) obtained for the surface meteorological sites and mobile measurements differed less than 1 K (1.5 K) for all months except January when RMSE (bias) was 4.8 K greater (5.1 K lower) for the mobile measurements. In January, WRF/Chem's performance in remote areas was clearly better than in the NAA. In combination with the MET–tower temperature gradients this finding indicates that in January WRF/Chem had difficulties with the strong inversions. In the NAA, sites are below the inversion, while most rural sites are in the mountains at altitudes above the inversion. Thus, those difficulties affected model performance more in than outside the NAA. Due to inversions, monthly averaged temperatures were between 1.2 and 5.9 K lower in the NAA than in rural areas except March. In March, melting began, and was faster in town leading to lower albedo and stronger heating of near surface air than in the mountains.

3.4. FRM vs. BAM

At the SB and PR, a FRM and BAM were run concurrently. The FRM and BAM 24 h–average $PM_{2.5}$ concentrations agreed within a factor of two for all but one value greater than 15 μg/m^3 at each site (Figure 3). The correlation between the BAM and FRM values was 0.73 (0.94) at the SB (PR). The BAM (FRM) measured minimum concentrations were 0 (1) at the SB, and 2.8 (1) at PR. The BAM (FRM) measured maximum concentrations were 135.3 μg/m^3 (140 μg/m^3) at the SB, and 114.1 μg/m^3 (82.5 μg/m^3) at PR. On average over OTM, at the SB, the FRM provided 3.3 μg/m^3 higher $PM_{2.5}$ concentrations, while at PR it provided 4.5 μg/m^3 lower $PM_{2.5}$ concentrations than the BAM. The mean error, FB, FE, NMB, and NME were 6 μg/m^3, –35%, 37%, –18% and 24%, respectively, at the SB, and 4.2 μg/m^3, 12%, 20%, 15% and 20% at PR site. Thus, we have to allow for uncertainty of this degree due to measurement errors in our evaluation. The relative high differences can be attributed to reduced accuracy at low temperature and relative humidity conditions.

3.5. $PM_{2.5}$ sites

We used all available model observation pairs and did not distinguish between observations from sites with hourly and 1–in–3–days sampling. Results from an analysis using only data from every third day marginally differed from those presented here.

Considering all sites over OTM, WRF/Chem failed to capture the tail of extremes to the fullest. It captured the frequency of concentrations best between 15 and 50 μg/m^3, and strongly

underestimated the frequency of concentrations less than 15 µg/m^3. It overestimated (underestimated) the frequency of concentrations around 70 µg/m^3 (greater than that). Around 35 µg/m^3 WRF/Chem tended to slight overestimation which explains the overestimation of the number of exceedances (Table 1).

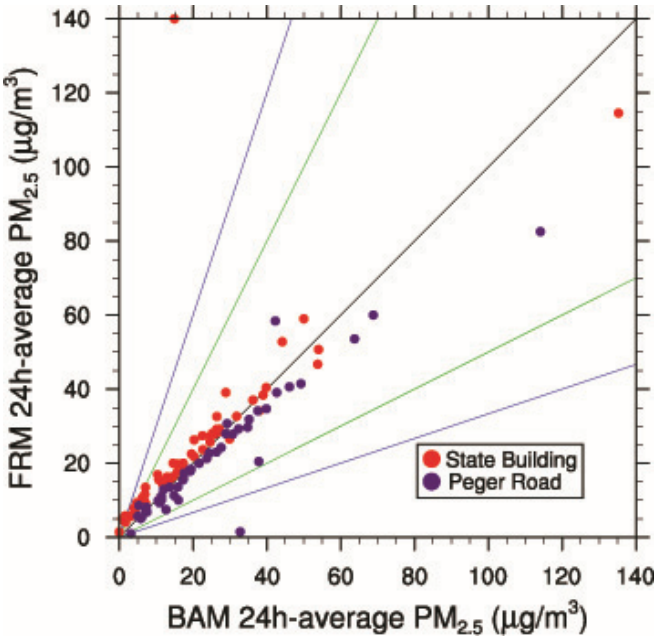

Figure 3. *Comparison of BAM and FRM–measured 24 h–average $PM_{2.5}$ concentrations at the SB and PR. Black, green and blue lines indicate the 1:1 line, factor of two and three agreement.*

Minimum simulated and observed $PM_{2.5}$ concentrations were 0.4 µg/m^3 at DP and 0 µg/m^3 at various sites. The highest hourly (179.6 µg/m^3) and 24 h–average (75.1 µg/m^3) $PM_{2.5}$ concentrations were simulated and observed (249.4 µg/m^3, and 135.3 µg/m^3, respectively) at the SB–BAM. The second highest hourly simulated (115.7 µg/m^3) and observed $PM_{2.5}$ concentrations (246.5 µg/m^3) occurred at the PR–BAM. The second highest 24 h–average $PM_{2.5}$ concentrations were simulated at Sadler/SB (75.1 µg/m^3) and observed at NP (116.7 µg/m^3). Obviously, WRF/Chem had difficulty capturing maxima in general.

Since Sadler and the SB fall into the same grid cell (Figure 1), WRF/Chem cannot capture the observed gradient between these sites that are 230 m apart. On average, WRF/Chem slightly better reproduced the Sadler than SB observations (Figure 4). This fact suggests a limited representativeness of the SB site for this grid cell. On average over OTM, WRF/Chem underestimated $PM_{2.5}$ concentrations at all sites except PR and NP.

In October, WRF/Chem failed to capture the temporal evolution of 24 h–average $PM_{2.5}$ concentrations at Sadler and the SB (Figure 4). This $PM_{2.5}$ overestimation may be caused by emissions that were allocated too high in the area in October. At the other sites except DP and RAM2, WRF/Chem captured the temporal evolution of 24 h–average $PM_{2.5}$ concentrations acceptably during OTM. The underestimation at DP on some days in February/March can be attributed to the use of background concentrations as boundary conditions. HYSPLIT backward meteorological trajectory simulations at 0000UTC and heights of 1 000 m to 8 500 m in steps of 500 m above ground for all days with concentrations higher than 2 µg/m^3 showed transport from Asia to DP at several levels (e.g. Figure 5).

Table 1. *$PM_{2.5}$ skill scores. Observed mean and standard deviation (StDev) are in brackets. No mobile measurements exist for October*

	October	November	December	January	February	March	1st quarter	4th quarter	OTM	November to March
24 h–averages at fixed locations										
Number of data	108	127	105	39	167	109	315	340	655	547
Mean (µg/m^3)	26.8(9.4)	20.0(20.8)	17.9(26.6)	32.0(40.5)	17.2(15.0)	11.2(7.7)	16.9(15.1)	21.0(19.6)	19.8(16.9)	17.7(17.8)
StDev (µg/m^3)	19.4(11.8)	11.2(18.4)	11.3(24.5)	19.3(23.1)	11.3(11.8)	6.7(5.4)	13.2(15.0)	14.7(20.3)	14.2(18.1)	12.5(18.7)
FB (%)	86	12	–28	–1	19	31	21	23	22	10
FE (%)	98	60	57	57	61	66	62	71	67	61
NMB (%)	167	–2	–34	–7	18	48	16	11	13	–3
NME (%)	181	60	53	58	64	77	65	76	71	60
Factor of two	35	63	57	59	61	60	60	52	56	60
number of exceedance days occurring at any of the sites in the NAA	19(3)	9(11)	5(14)	10(7)	10(9)	0(0)	20(16)	33(28)	53(44)	34(41)
1 h–averages at fixed locations										
Number of data	1 792	2 086	1 509	1 421	1 326	1 431	4 178	5 387	9 565	7 773
Mean (µg/m^3)	28.3(11.7)	23.3(22.1)	21.6(31.2)	26.3(22.3)	20.6(18.9)	13.2(8.9)	20.0(16.5)	24.5(21.2)	22.5(19.2)	21.2(20.9)
StDev (µg/m^3)	25.6(14.6)	14.6(19.9)	15.9(31.5)	22.1(22.2)	15.5(17.7)	9.2(10.6)	17.3(18.4)	19.5(23.6)	18.7(21.7)	16.4(22.6)
FB (%)	87	26	–8	28	14	56	33	37	35	23
FE (%)	113	91	85	91	92	113	99	97	98	94
NMB (%)	142	5	–31	18	9	53	21	16	18	1
NME (%)	176	83	70	85	87	132	94	95	94	84
Factor of two	33	41	44	43	40	30	38	39	39	40
Mobile measurements										
Mean (µg/m^3)		22.4(26.1)	18.7(25.1)	24.9(35.9)	26.5(9.32)	15.1(5.3)	24.2(24.8)	19.7(25.3)		22.2(25.0)
StDev (µg/m^3)		14.4(22.2)	15.5(28.7)	19.2(37.2)	16.4(9.9)	11.5(9.3)	18.0(32.5)	15.3(27.2)		16.9(30.2)
FB (%)		10	13	10	103	121	50	12		32
FE (%)		79	95	102	118	134	110	91		101
NMB (%)		–12	–26	–31	185	184	–2	–22		–11
NME (%)		63	78	79	211	238	97	74		87
Factor of two		52	36	31	25	13	27	40		33

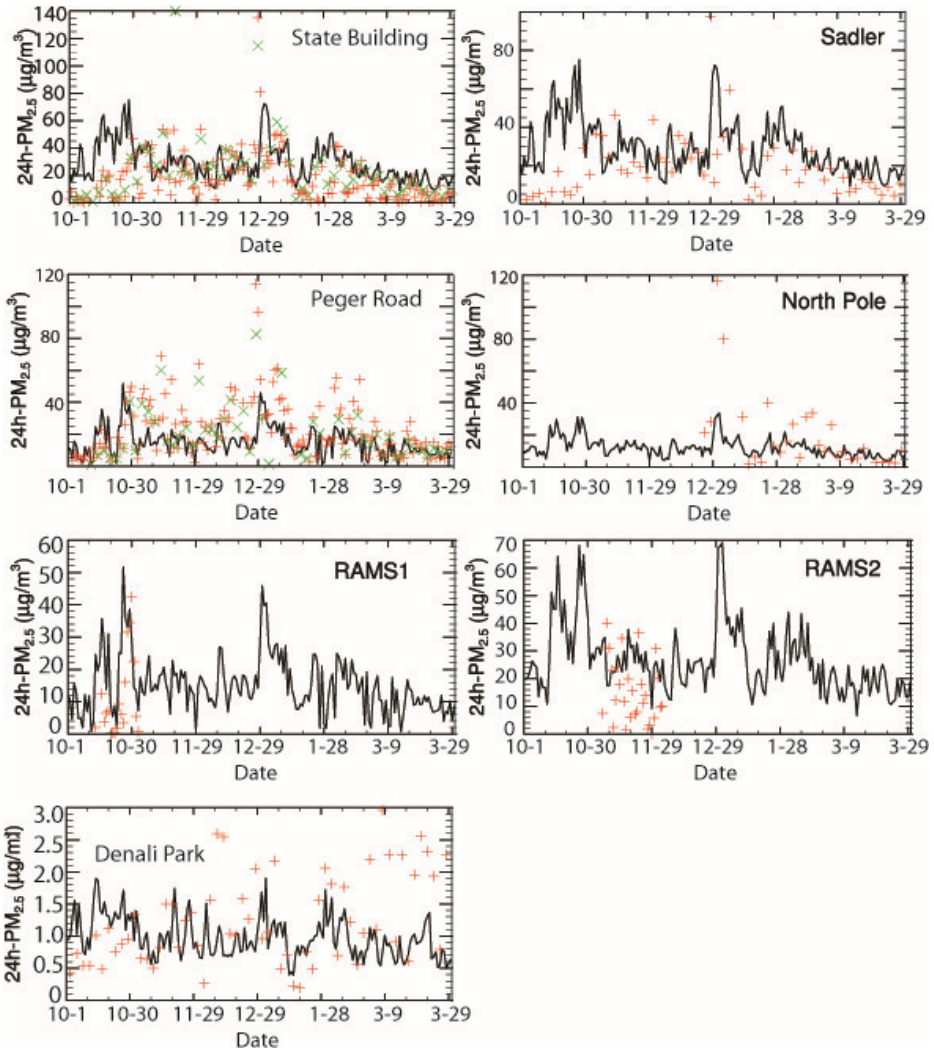

Figure 4. *Time series of simulated (line) and observed (symbols) $PM_{2.5}$ concentrations at various sites. For the SB and PR red (green) symbols indicate the BAM (FRM) data.*

At NP and PR, WRF/Chem failed to capture the full magnitude of concentration peaks (Figure 4). Errors in emissions and meteorology are most likely the causes. Both sites are in flat terrain, but at the NP site, residential heating emissions instead of traffic and industrial sources dominate the emissions. Emissions from traffic and heating were allocated temperature dependent, and at both sites, the largest discrepancies occurred in January and February when the temperature biases were 2.6 and 1.6 K domain–wide and 1.7 and 0.3 K in the NAA.

On average, WRF/Chem captured the temporal evolution of $PM_{2.5}$ concentrations acceptably except during sudden temperature changes, underestimated inversion strengths and mistimed frontal passages (Figures 2 and 4). Discrepancies in the hourly temporal evolution may result from channeling effects in streets and/or slight offsets of dispersion plumes due to errors in wind direction. The occasionally much higher observed than simulated concentrations are most likely due to contamination of the measurements by mobile sources (At temperatures below –25 °C, many Fairbanksans idle their parked cars and most sites are in parking lots.). Discrepancies due to such issues are common in all AQMs of the scale deployed here (Chang and Hanna, 2004).

Over OTM for the seven $PM_{2.5}$ sites, WRF/Chem met the performance criteria and goals at six and one site, respectively (Figure 6). It met the combined performance criteria at 2, 5, 7, 5, 5 and 4 sites, and the combined performance goals at 1, 4, 3, 2, 2, and 2 sites from October to March, respectively. In the first (fourth) quarter, WRF/Chem met the performance goals at 2 (1) and criteria at 2 (4) sites. At PR (both instruments) and NP, WRF/Chem met the performance criteria in all months. At the SB–FRM, it met them in all months except October and March (Figure 6). Performance was best for high average concentrations (>30 µg/m^3) like in December, and worst for October and average concentrations around 10 µg/m^3. At DP, forecasts were of similar quality on weekends and weekdays. Forecasts were better on weekdays for the SB–BAM, Sadler, and both RAMS sites, while the other sites' forecasts were better on weekends. The different performance on weekends than weekdays in the NAA hints at inaccurate emissions as major contributor for forecast errors.

The scientific community considers AQMs with FB within ±30%, and 50% of the forecasts falling within a factor of two of the observations as good (Chang and Hanna, 2004). Over all sites, FB was below 30% for the 24 h–average and hourly $PM_{2.5}$ except October and March (Table 1). Typically, WRF/Chem performed better for the 24 h–average than hourly $PM_{2.5}$ concentrations as slight offsets due to meteorology (Figures 2 and 4) or emissions can cancel out over 24 h. Over all sites and OTM, the FB of 24 h–average $PM_{2.5}$ concentrations was 22%, and ranged between –27% and 70% at all sites except RAMS1. Over OTM 56% of the simulated and observed 24 h–average and 39% of the hourly $PM_{2.5}$ concentrations agreed within a factor of two. Best (worst) agreements with 49% (21%) and 73% (26%) occurred for the hourly values at PR (SB) and the 24 h–averages at DP (RAMS1). In the NAA, forecasts and observations agreed best within a factor of two (67%) at PR.

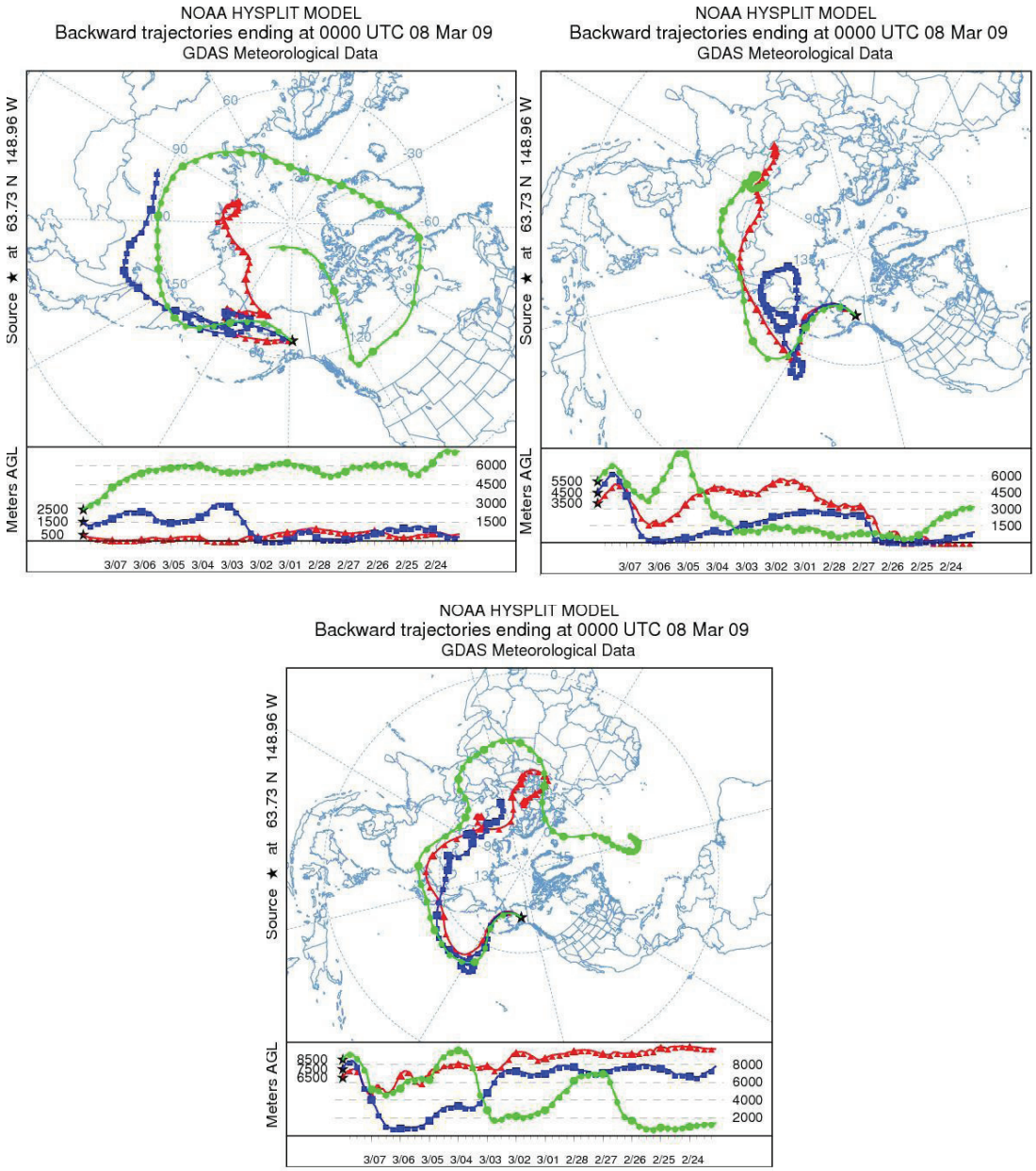

Figure 5. *HYSPLIT backward meteorological trajectory simulations for March 8 2009 from DP at three heights for each simulation.*

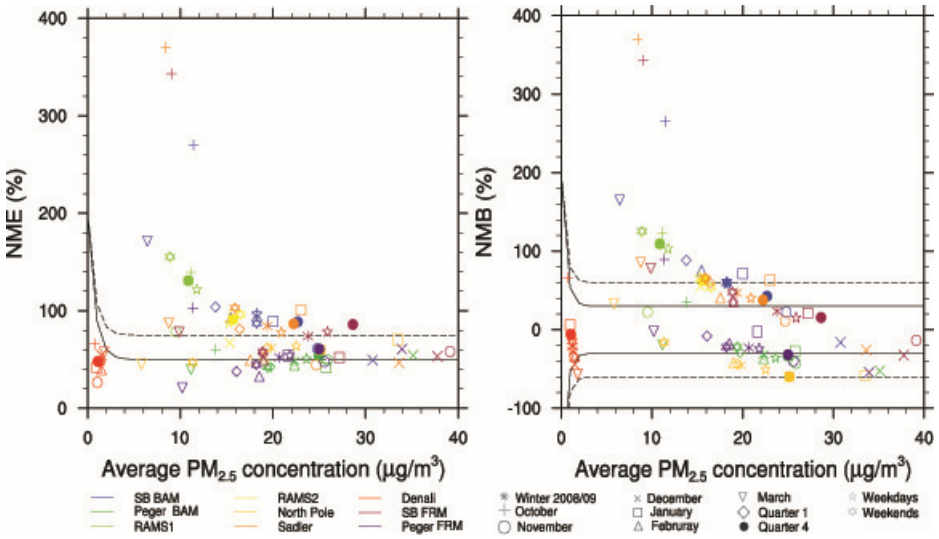

Figure 6. *NME and NMB as function of average concentration. The solid (dashed) lines indicate the performance goals (criteria).*

$PM_{2.5}$ concentrations were most likely simulated correctly when temperature was simulated correctly (Figure 7). Frequently temperature overestimation of about 6 K still led to correct $PM_{2.5}$ concentrations. This phenomenon may be partly attributed to the temperature effect on the nucleation process and sometimes to low concentrations. As long as the critical value for nucleation, which depends on temperature, is not exceeded no particles form.

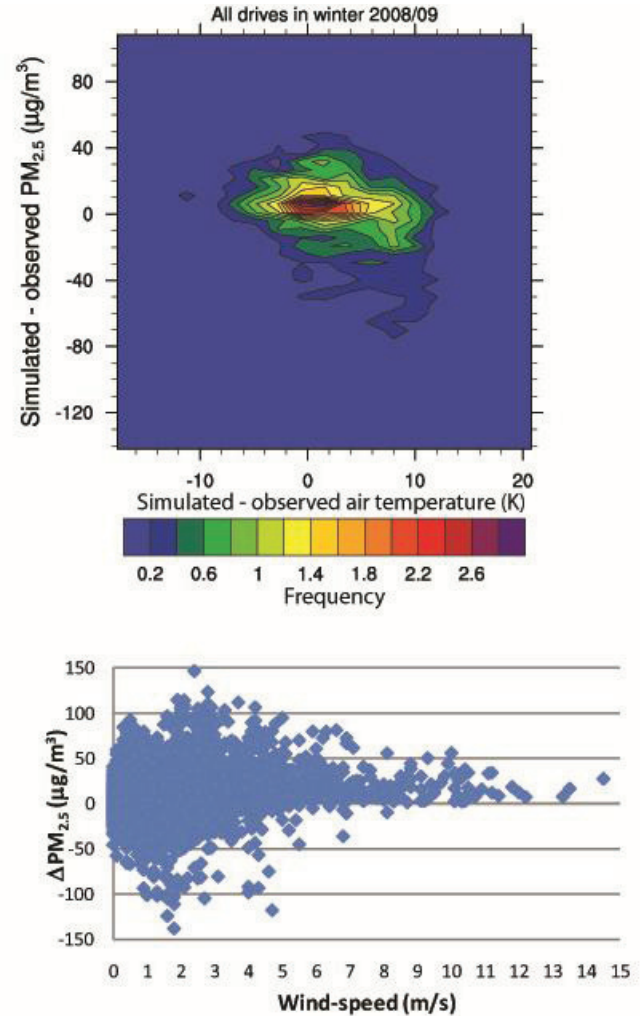

Figure 7. *Frequency of temperature error and $PM_{2.5}$ error pairs for OTM and all sites* ***(upper)****, and $PM_{2.5}$ errors at the SB as a function of wind speed at the MET–tower* ***(lower panel)****.*

Some of the $PM_{2.5}$ underestimation related to the on average overestimated relative humidity. Since under subarctic conditions, particles swell when relative humidity exceeds 70% (Tran and Mölders, 2011), the too high relative humidity shifts $PM_{2.5}$ towards PM_{10}. Our evaluation suggested that under low irradiation subarctic conditions (November to January) WRF/Chem may convert $PM_{2.5}$ too quickly to PM_{10}. Based on the IMPROVE data, WRF/Chem captured the $PM_{2.5}$ perturbations in pristine air on average well except during events with advection from Asia (Figures 4 and 5).

3.6 $PM_{2.5}$ speciation

The speciation data revealed better performance for relatively high concentrations of the $PM_{2.5}$ components (Figure 8). Performance differed with species and site. Simulated and observed speciation agreed well at all sites for SO_4^{2-}, OC and EC (Table 2). Time series of skill scores indicated no trends, only slightly varying performance. The skill scores did not correlate statistically significantly with errors in simulated meteorology (therefore not shown). This suggests inaccurate emissions as an error source. The different performance among sites for the same species supports this suggestion. The "polluted" speciation sites are in three different neighborhoods, although all below the inversion layer. The SB and PR sites see strong traffic emissions, while NP and SB also see strong contributions from heating. At the remote IMPROVE site, errors in meteorology and chemical boundary conditions are the most likely cause.

Table 2. *Speciation skill scores based on 40, 15, 15 and 59 days with data at the SB, PR, NP, and DP–sites. NH_4^+–FBs, FEs, NMBs, NMEs and factor of two are –199%, 199%, –100%, 100%, and 0%, respectively, at all sites except DP where no NH_4^+ and NO_3^- data were available*

		FB (%)	FE (%)	NMB (%)	NME (%)	Factor of two (%)
OC	SB	18	43	11	46	78
	PR	–64	69	–47	50	47
	NP	–64	72	–59	63	47
	DP	81	87	110	119	34
EC	SB	55	65	63	79	53
	PR	–37	53	–29	39	80
	NP	19	58	0	56	60
	DP	100	106	154	169	31
SO_4^{2-}	SB	40	61	29	70	60
	PR	–13	56	–11	52	80
	NP	37	41	43	50	87
	DP	–6	55	–24	50	59
NO_3^-	SB	–125	129	–71	76	25
	PR	–164	164	–90	–90	0
	NP	–165	165	–89	–89	0

Organic carbon is primarily emitted or formed from the condensation of low volatility hydrocarbons. Thus, OC performance depends on the accuracy of emissions and simulated gas–to–particle conversion. In Fairbanks, major primary OC sources are wood burning, diesel and gasoline engines, and some industrial processes. OC forecasts were best at the SB. On average over OTM (Figure 8), simulated and observed OC agreed well at the SB with 78% being within a factor of two (Table 2). Here, WRF/Chem met the combined performance goals in all months except February and March. In NP and at PR, WRF/Chem underestimated OC most time. At NP, performance was weakest in February, but still 30% (70%) of the simulated and observed OC agreed within a factor of two (three). At DP, WRF/Chem overestimated OC except in March (Figure 8).

EC performance was better at NP and PR than at the SB where still 53% of the simulated EC values agreed within a factor of two with the observations (Table 2). At the SB, WRF/Chem met the performance goal in December, and the performance criteria in the fourth quarter, OTM, and November (Figure 8). At NP and DP, it met the performance criteria in all months. At PR, WRF/Chem met the performance goals (criteria) in February, the first quarter, OTM and on weekdays (weekends, January), while it failed the performance criterion in March. Obviously, the slightly lower EC concentrations in the first quarter affected that quarter's EC performance.

The on average overestimation of OC and EC at the remote and underestimation at the polluted NP and PR sites suggest that the on average overestimated wind speed (Figure 2) over–diluted the pollutants.

MADE/SORGAM considers, among others, NH_4^+, SO_4^{2-}, NO_3^-, water and other inorganic aerosol components. Sulfate forms from SO_2 by gas–to–particle conversion or aqueous–phase reactions. Once formed, it reacts with NH_3 to ammonium sulfate ($(NH_4)_2SO_4$).

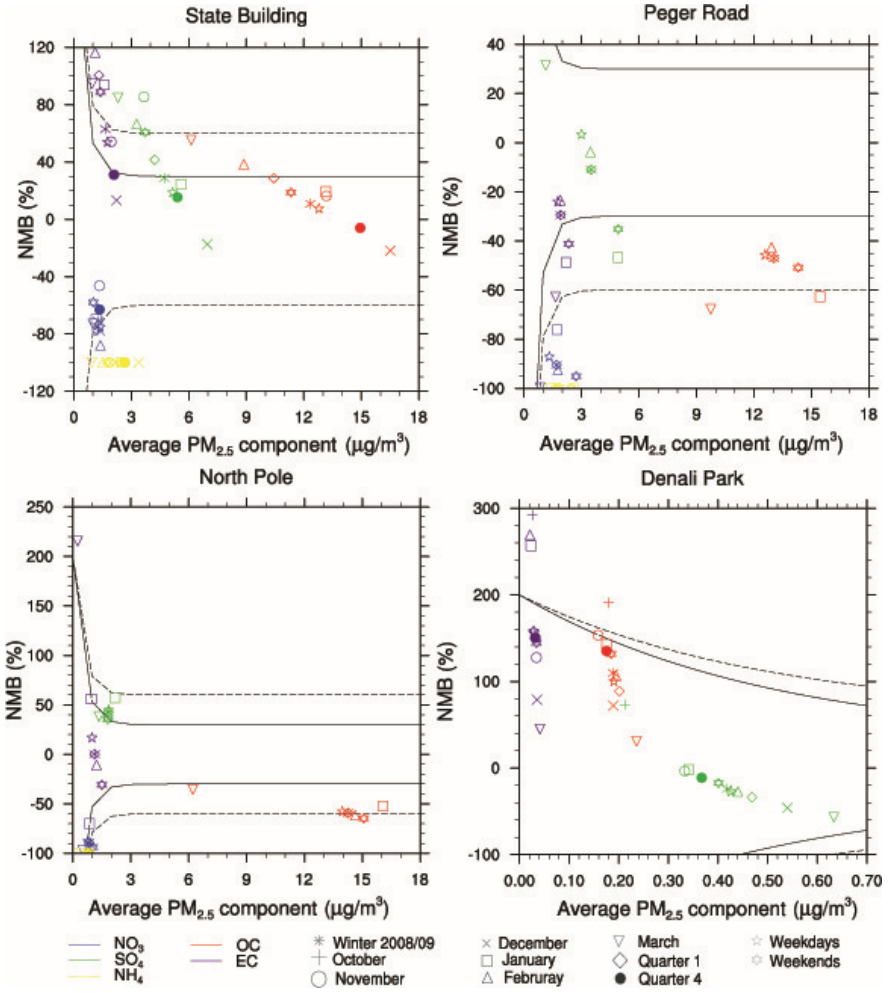

Figure 8. *NMB as function of average speciation concentration. The solid (dashed) lines indicate the performance goals (criteria).*

Ammonium forms as NH_3 neutralizes H_2SO_4 and HNO_3. The amount of NH_4^+ depends highly on the relative amounts of H_2SO_4 and NH_3. Nitric acid can react with the excess NH_3 to ammonium nitrate (NH_4NO_3). Given these two potential paths, two NH_3 molecules yield either $2NH_4NO_3$ or $1(NH_4)_2SO_4$.

Sulfate concentrations were lower in NP than at the SB and PR that had similar magnitudes. Simulated and observed SO_4^{2-} agreed within a factor of two, 59% or more most of the time (Table 2). At the SB, WRF/Chem met the performance criteria for SO_4^{2-} in October, December and January (Figure 8). Then on average, slightly higher SO_4^{2-} concentrations yielded better performance than in November, February and March. At NP and PR, WRF/Chem met the combined performance criteria for all months, the first quarter, weekends and weekdays. At the SB, WRF/Chem simulated SO_4^{2-} better on weekdays than weekends, while the opposite was true for NP. At DP, WRF/Chem underestimated SO_4^{2-} except in October and met the performance criteria in all months.

Nitrate is difficult to simulate due to its high volatility and sensitivity to NH_3 availability, temperature and relative humidity. At the SB and NP sites, the monthly, weekend, and weekday average NO_3^- concentrations differed marginally, while at PR, they varied notably among January, March, weekdays and weekends (Figure 8). As temperature and humidity did not vary substantially over town, the performance differences may be attributed to the different emissions in these neighborhoods. WRF/Chem captured NO_3^- best at the SB. Here 25% of the simulated and observed NO_3^- concentrations agreed within a factor of two (Table 2). At the SB, WRF/Chem met the performance criteria in November and on weekends. At the other sites, it never met these criteria, even though it only marginally missed them for NP.

At all sites, monthly averaged observed NH_4^+ was less than 4 µg/m^3. Compared to other species, NH_4^+ performance was weak at all sites (Table 2). On average over OTM, none of the simulated NH_4^+ values agreed within a factor of two with the observations at any site. The too low NH_4^+ concentrations explain some of the discrepancies between simulated and observed SO_4^{2-} andNO_3^-.

The NEI2008 NH_3 emissions showed that there is hardly any agriculture and livestock. The strong NH_4^+ underestimation at all sites (Figure 8), however, suggested that some NH_3 emission sources are overlooked. Fairbanks has a huge number of dog kennels with more than 30 animals each who may contribute to NH_3 emissions. Lee and Dollard (1994) estimated that in the UK, NH_3 emissions from pets may be of about 1/6th the magnitude of the NH_3 emission from agriculture.

3.7. Mobile $PM_{2.5}$ measurements

Model performance depends on the representativeness of the measurements (Park et al., 2006; PaiMazumder and Mölders, 2009). The mobile measured concentrations were taken on roads with traffic. Commonly, it is assumed that measurements from well placed sites are representative for a radius of 4 km or so (Park et al., 2006). To examine the representativeness of the data we compared the spatially averaged mobile measurements taken within a 1, 2 and 4 km radius of fixed sites with the concentrations

measured at these sites when the vehicle was in their vicinity. Four sites (PR, SB, RAMS1, and RAMS2) were within these distances of the vehicles routes. Differences between area averaged concentrations and concentrations at the sites did not change notably with increasing radius (therefore not shown). This means the use of hourly spatially averaged observations and the dense network of routes yielded small gradients.

The mobile measurements indicated $PM_{2.5}$ concentrations to be relatively high in the southwestern NAA and around North Pole, extremely low east of North Pole, increased close to Moose Creek and lower towards Eielson. On average over all drives, WRF/Chem provided slightly lower (2.8 μg/m^3) $PM_{2.5}$ concentrations than the mobile measurements (Table 1). For November and January, using the mobile data provided slightly lower FB than using the hourly data of fixed sites. However, the agreement between simulated and observed values is less than for the fixed sites 1 h–data except November and the 24 h–average data. In February and March, WRF/Chem overestimated the concentrations strongly in the grid cells with mobile measurements, but less at the fixed sites. This finding supports the conclusion that WRF/Chem has difficulties to capture the local maxima and minima.

The differences in performance found with the different $PM_{2.5}$ datasets (Table 1) result from the fact that the fixed location data provide the temporal evolution with low spatial resolution (low number of sites), while the mobile measurements provide high spatial resolution, but only for a couple of hours on days with drives. Moreover, mobile measurements were taken mostly on days with high pollution at PR and/or the SB. The differences suggest that WRF/Chem has difficulties capturing local maxima and minima.

3.8. Discussion

Under similar temperature conditions weekend and weekday performance differed at the various sites. This fact suggests that emissions contributed to the discrepancies. According to both the mobile and surface meteorological measurements, temperature forecasts were worst for November, and wind forecasts were worst in October and February (Figure 2). However, $PM_{2.5}$ forecasts were overall weaker for October, February and March than for the other months (Table 1). These findings hint at incorrect emissions as a strong contributor for discrepancies in March and at wind speed errors in February and October. The obvious different performance among sites in October suggests locally incorrect emissions as a further cause. In October and March, temperature performance was best. In January, relative humidity and wind speed were predicted best. Thus, discrepancies were mainly due to incorrect emissions and temperature.

Some of the relatively lower performance compared to mid–latitude applications is related to the low wind speeds in Interior Alaska. Wind speeds less than 1 m/s dominated from November through February, and were on average less than 2 m/s in October and March. During OTM, simulated (observed) wind speed was less than 1 m/s for 91% (34)% of the time at the MET–tower. WRF/Chem seldom underestimated $PM_{2.5}$ concentrations when observed wind speeds exceeded 4 m/s (Figure 7). The wind speed and direction errors affect the pollution dilution and direction of the simulated plume with consequences for simulated $PM_{2.5}$ concentrations.

Compared with the 4 km–increment MM5–CMAQ January simulations performed for North Carolina by Liu et al. (2010) our $PM_{2.5}$ forecasts are of slightly weaker quality, on average. This lower performance may be partly explained by the larger errors in meteorological quantities in our January simulations as compared to their's. However, our performance skills for reproducing the meteorological variables were consistent with current model performance for Alaska winter simulations (Mölders et al., 2011). $PM_{2.5}$ forecasts were most frequently within ±5 μg/m^3 when temperature errors were within –2 K to 6 K of the observations (Figure 7).

Our evaluation of simulated 24 h–average $PM_{2.5}$ concen–trations (Table 1) provided slightly higher NMB and NME than other WRF/Chem–studies. Zhang et al.'s (2010) 36 km–increment January simulations over the continental US had NMB of 32.2 and –6.7% and NME of 70.9 and 63.9% when evaluated by the IMPROVE and STN–data, respectively. Our January–NMB and NME were –7% and 58% for the fixed location sites. The grid increment may affect skill scores. Misenis and Zhang (2010) reported that using a 4 km–increment improved the $PM_{2.5}$ forecasts compared with their 12 km–increment–simulation. Thus, even though our NMB and NME only slightly differ from those of Zhang et al.' 36 km–increment–simulations the actual performance for high–latitudes may be slightly lower than for mid–latitudes.

The low data density and $PM_{2.5}$ site distribution provide challenges for model improvement for high–latitude applications. Liu et al. (2010) had 836 observations for their North Carolina $PM_{2.5}$ evaluation for January. They reported spatial differences in performance and NMBs between 8.5 and 17.1%, and NMEs between 27.6 and 38.5% for their 4 km–increment MM5/CMAQ January simulation using data from three networks. In our study in January, only 39 $PM_{2.5}$ observations, spread among seven sites, existed where all but one $PM_{2.5}$ site was in the NAA. Thus, when the model performed weakly in the NAA, but well in other areas this spatial inhomogeneity remains undetected. The scarce data in high–latitudes reduces the opportunity to examine reasons for better performance in one area than in another, which would help to improve the model. Thus, the network needs to be extended into the remote areas to assess the performance for high–latitude conditions fully and improve AQMs for high–latitude applications. Having measurements aloft would provide insight in how difficulties in capturing the inversion strength affect the $PM_{2.5}$ vertical profiles.

4. Conclusions

We used mobile and fixed location measurements from the Fairbanks winter 2008/09 field campaign for an operational evaluation of WRF/Chem's performance at high–latitudes. Based on the skill scores determined over all sites, WRF/Chem simulated $PM_{2.5}$ concentrations well in OTM. WRF/Chem performed best for $PM_{2.5}$ concentrations between 15 and 50 μg/m^3, but failed to capture the extremes (low, high concentrations). Since the NAAQS and the design value for 2008 for the Fairbanks NAA of 44.7 μg/m^3 fall into the range of best performance, we may conclude that WRF/Chem is suitable for examining the impact of emission reduction scenarios.

At all speciation sites, the performance accuracy of speciation increased with species concentration and was higher for the more abundant species. However, at sites with low maximum concentration, performance may exceed that at sites with high maximum concentrations of that species. This finding hints at errors in local emissions. WRF/Chem performed better for EC, OC and SO_4^{2-} than for NO_3^- and NH_4^+. Nevertheless, since WRF/Chem explicitly calculated secondary organic aerosol, the acceptable agreement between pairs of simulated and observed species suggests that WRF/Chem is able to reasonably reproduce the sensitivity of secondary organic aerosol to emission control strategies.

Our evaluation provided a benchmark of WRF/Chem's performance in a high–latitude environment and identified the following uncertainties and limitations. The comparison of measurements taken concurrently at the same site with two different devices suggested, on average over OTM, an observational uncertainty in mean bias, mean error, FB, FE, NMB,

and NME of 4.5 μg/m^3, 5.8 μg/m^3, –35%, 37%, –18% to 15% and 24%, respectively.

Measurement errors due to the extreme low temperature and moisture conditions are not the major cause for the lower performance than in mid–latitude applications. Other error sources impacting model performance included the emission inventory, simulated meteorology, parameters in chemical mechanisms, low representativeness of some sites, initial and boundary conditions. The speciation data indicated that the NH_3 emissions are not fully understood/known. The time series of simulated and observed $PM_{2.5}$ indicated inaccuracy in emissions near at least two sites in October.

The on average, over all sites and OTM, SLP, air and dewpoint temperatures, wind speed and direction biases of –1.9 hPa, 1.3 K, 2.1 K, 1.55 m/s, and –4°, respectively, contributed to $PM_{2.5}$ forecast errors in various ways. Performance in simulating temperature was better at relatively higher temperatures and in rural areas. The underestimation of the inversion strength and overestimation of wind speed also contributed to concentration underestimations.

The study suggested that in the NAA, the representativeness of sites can be limited. The availability of mobile measured $PM_{2.5}$ data allowed assessing WRF/Chem's ability to capture the horizontal heterogeneity of concentrations. WRF/Chem had difficulties in capturing the local maxima and minima.

Comparison with other studies identified that the network has to be extended to improve/further develop AQMs for high–latitude applications. Low data density means that a few incorrect measurements and/or poorly located sites affect the apparent performance much higher than they would in regions with high, area–wide site–density.

Acknowledgements

We thank G. Kramm and the anonymous reviewers for fruitful discussion, J. Conner, J. McCormick, D. Huff, R. Lovel and N. Svengards for the observational data, the Fairbanks North Star Borough (contract LGFEEQ), Department of Transportation (contract AUTC410003) and NSF (contracts ATM0630506, cooperative agreement #0919608) for financial support.

List of acronyms

AQM	air quality model
BAM	Beta Attenuation Monitors
DP	Denali Park
FB	mean fractional bias
FE	mean fractional error
FRM	Federal Reference Method
IMPROVE	Interagency Monitoring of Protected Visual Environments
NP	North Pole
MADE	Modal Aerosol Dynamics Model for Europe
NAA	nonattainment area
NAAQS	National Ambient Air Quality Standard
NEI	National Emission Inventory
NMB	normalized mean bias
NME	normalized mean error
OC	organic carbon
OTM	October to March (10–1–2008 to 31–3–2009)
PR	Peger Road
QA/QC	quality assurance/quality control
RAMS	Regional Air Monitoring System
RMSE	root mean square error
SB	State Office Building
SLP	sea–level pressure
SORGAM	Secondary Organic Aerosol Model
STN	Speciation Trends Network
WRF/Chem	Weather Research and Forecasting model with inline chemistry package

References

Ackermann, I.J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F.S., Shankar, U., 1998. Modal aerosol dynamics model for Europe: development and first applications. *Atmospheric Environment* 32, 2981-2999.

Boylan, J.W., Russell, A.G., 2006. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmospheric Environment* 40, 4946-4959.

Cahill, C.F., 2003. Asian aerosol transport to Alaska during ACE-Asia. *Journal Geophysical Research* 108, art. no. 8664.

Chang, J.C., Hanna, S.R., 2004. Air quality model performance evaluation. *Meteorology and Atmospheric Physics* 87, 167-196.

Djalalova, I., Wilczak, J., McKeen, S., Grell, G., Peckham, S., Pagowski, M., DelleMonache, L., McQueen, J., Tang, Y., Lee, P., McHenry, J., Gong, W., Bouchet, V., Mathur, R., 2010. Ensemble and bias-correction techniques for air quality model forecasts of surface O_3 and $PM_{2.5}$ during the TEXAQS-II experiment of 2006. *Atmospheric Environment* 44, 455-467.

Draxler, R., Stunder, B., Rolph, G., Stein, A., Taylor, A., 2009. HYSPLIT4 User's Guide. Version 4.9.

Grell, G.A., Devenyi, D., 2002. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters* 29, art. no. 1693.

Hong, S.-Y., Lim, J.-O.J., 2006. The WRF single-moment 6-class microphysics scheme (WSM6). *Journal of the Korean Meteorological Society* 42, 129-151.

Janjic, Z.I., 2002. Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso Model. National Centers for Environmental Prediction (NCEP) Office Note #437, 61 pp.

Lee, D.S., Dollard, G.J., 1994. Uncertainties in current estimates of emissions of ammonia in the United Kingdom. *Environmental Pollution* 86, 267-277.

Liu, X.H., Zhang, Y., Olsen, K.M., Wang, W.X., Do, B.A., Bridgers, G.M., 2010. Responses of future air quality to emission controls over North Carolina, Part I: model evaluation for current year simulations. *Atmospheric Environment* 44, 2443-2456.

Madronich, S., 1987. Photodissociation in the atmosphere. 1. actinic flux and the effects of ground reflections and clouds. *Journal of Geophysical Research-Atmospheres* 92, 9740-9752.

Misenis, C., Zhang, Y., 2010. An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options. *Atmospheric Research* 97, 315-334.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research-Atmospheres* 102, 16663-16682.

Mölders, N., Tran, H.N.Q., Quinn, P., Sassen, K., Shaw, G.E., Kramm, G., 2011. Assessment of WRF/Chem to capture sub-Arctic boundary layer characteristics during low solar irradiation using radiosonde, SODAR, and surface data. *Atmospheric Pollution Research* 2, 283-299.

PaiMazumder, D., Mölders, N., 2009. Theoretical assessment of uncertainty in regional averages due to network density and design. *Journal of Applied Meteorology and Climatology* 48, 1643–1666.

Park, S.K., Cobb, C.E., Wade, K., Mulholland, J., Hu, Y.T., Russell, A.G., 2006. Uncertainty in air quality model evaluation for particulate matter due to spatial variations in pollutant concentrations. *Atmospheric Environment* 40, S563-S573.

Peckham, S.E., Fast, J.D., Schmitz, R., Grell, G.A., Gustafson, W.I., McKeen, S.A., Ghan, S.J., Zaveri, R., Easter, R.C., Barnard, J., Chapman, E., Salzmann, M., Wiedinmyer, C., Freitas, S.R., 2009. WRF/Chem Version 3.1 User's Guide, 78 pp.

Schell, B., Ackermann, I.J., Hass, H., Binkowski, F.S., Ebel, A., 2001. Modeling the formation of secondary organic aerosol within a comprehensive air quality model system. *Journal of Geophysical Research-Atmospheres* 106, 28275-28293.

Simpson, D., Guenther, A., Hewitt, C.N., Steinbrecher, R., 1995. Biogenic emissions in Europe. 1. estimates and uncertainties. *Journal of Geophysical Research-Atmospheres* 100, 22875-22890.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., Powers, J.G., 2008. A description of the Advanced Research WRF Version 3, NCAR Technical Note, NCAR/TN–475+STR, Colorado, USA, 125 pp.

Smirnova, T.G., Brown, J.M., Benjamin, S.G., Kim, D., 2000. Parameterization of cold-season processes in the MAPS land-surface scheme. *Journal of Geophysical Research-Atmospheres* 105, 4077-4086.

Stockwell, W.R., Middleton, P., Chang, J.S., Tang, X., 1990. The second generation regional acid deposition model chemical mechanism for regional air quality modeling. *Journal of Geophysical Research-Atmospheres* 95, 16343-16367.

Tran, H.N.Q., Mölders, N., 2011. Investigations on meteorological conditions for elevated $PM_{2.5}$ in Fairbanks, Alaska. *Atmospheric Research* 99, 39-49.

Tran, T.T., Newby, G., Mölders, N., 2011. Impacts of emission changes on sulfate aerosols in Alaska. *Atmospheric Environment* 45, 3078-3090.

Wesely, M.L., 1989. Parameterization of surface resistances to gaseous dry deposition in regional scale numerical models. *Atmospheric Environment* 23, 1293-1304.

Zhang, Y., Wen, X.Y., Jang, C.J., 2010. Simulating chemistry-aerosol-cloud-radiation-climate feedbacks over the continental US using the online-coupled weather research forecasting model with chemistry (WRF/Chem). *Atmospheric Environment* 44, 3568-3582.

Zhao, Z., Chen, S.H., Kleeman, M.J., Tyree, M., Cayan, D., 2011. The impact of climate change on air quality-related meteorological conditions in California. Part I: present time simulation analysis. *Journal of Climate* 24, 3344-3361.